Automatic recognition of emotional dimensions in singing

Florian Eyben[1], Eduardo Coutinho[2]

1: AUDeering, Germany; 2: University of Liverpool, UK

In this contribution we evaluate the feasibility of automatic recognition of emotion portrayals in singing, and the importance of devising a comprehensive set of relevant acoustic parameters. Computational experiments and validation are performed on the Geneva Corpus of Emotional Singing (GeCESi). The emotional states are mapped to ternary arousal and valence labels for simplified interpretation of the results. A small set of relevant acoustic features derived from our previous research on the same data and a large scale, state-of-the-art feature set commonly used for the recognition of paralinguistic phenomena (the baseline feature set of the Interspeech 2013 Computational Paralinguistics ChallengE; ComParE) were used. A feature relevance analysis with respect to classification accuracy and correlation of features with the targets is conducted and discussed. The results show that the automatic classification performance for arousal is similar for both feature sets, while the ComParE set is superior for valence. It is further shown that applying a novel intra singer feature ranking criteria to select robust features improves the generalisation of the automated classification to new singers even further.